DOCUMENT RESUME

ED 103 831 CS 001 695

AUTHOR Allington, Richard L.
TITLE A Comparison of Taped Micro-Teaching Performance and Teacher Test Achievement.
PUB DATE Nov 74
NOTE 17p.; Paper presented at the Annual Meeting of the Northeastern Educational Research Association (5th, Ellenville, New York, November 1974)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS College Students; Comparative Analysis; Evaluation Methods; Graduate Study; Higher Education; *Microteaching; *Multiple Choice Tests; *Reading Instruction; *Reading Research; *Teaching Methods; Test Results

ABSTRACT

This study investigated the relationship between instructor ratings of taped microteaching performance and achievement on multiple choice examinations. Thirty graduate students enrolled in a reading methods course served as subjects. An analysis of the data indicated that no significant correlation existed between the variables. The viability of the taped minilessons as an alternative to the traditional paper-pencil assessment techniques is discussed. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED103831

A Comparison of Taped Micro-Teaching Performance and Teacher Test Achievement

Richard L. Allington, Ph.D.
Assistant Professor
Reading Department
State University of New York
at Albany
Albany, New York 12222

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Richard L. Allington
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Paper presented at the 5th Annual Convocation of the Northeastern Educational Research Association, November 1, 1974 (Ellenville, N.Y.).

CS 001 695

Abstract

This study investigated the relationship between instructor ratings of taped micro-teaching performance and achievement on multiple choice examinations. Thirty graduate students enrolled in a readings methods course served as subjects. An analysis of the data indicated no significant correlation existed between the variables. The viability of the taped mini-lessons as an alternative to the traditional paper-pencil assessment techniques is discussed.

A Comparison of Taped Micro-Teaching Performance and Teacher Test Achievement

Teacher education has suffered recurrent criticism of its seeming inability to ensure teaching proficiency in its graduates. A major issue has been the assessment of teaching performance versus the assessment of cognitive knowledge (Allington, 1974). The advent of competency-based teacher education has spawned experimentation with a variety of techniques and instruments for predicting teaching performance in an attempt to rectify this weakness. However, teacher trainers are caught in a continual struggle between assessment efficiency and assessment effectiveness. While efficient, the widely used and traditional paper-pencil examination seems less than effective in predicting teaching performance. Shea (1955), for example, administered the National Teachers Examination (NTE) to 110 graduates of a four year teacher training institution. The correlations between practice teaching grades and the NTE yielded a correlation of .00. Following a factor analysis of the data Shea concluded, "that none of the standardized tests, with the possible exception of subtests that measure general information..., are factorially related to marks obtained...on practice teaching, (or) on the job rating (p. 36)." Walberg (1967) drew similar conclusions in a study of 280 elementary teachers. His data demonstrated no correlation between teaching effectiveness and any of the measures of academic aptitude and achievement employed (high school grades, college

aptitude tests, college grade point average, or National Teacher Examination score). Similar conclusions have been drawn by a number of researchers and summaries of these can be found in reviews by Barr (1961), Getzels and Jackson (1963) and Levin (1968). While efficient, traditional paper-pencil assessments seem to offer limited promise for predicting teaching performance.

Research results such as these were the impetus for the development of alternative assessment techniques. In many teacher education programs the logical alternative, direct field assessment, has been rejected primarily out of cost considerations. How does one then establish that competence does exist?

Popham (1971) has discussed the positive effects of mini-lessons. Simulation teaching and role-playing have also been utilized as assessment or instructional techniques (Dodl, 1973). These techniques seem to have developed because of the aforementioned difficulty in arranging (or financing) field assessment of teaching competence.

In the preparation of reading teachers one might include many competencies, but competence in identifying skills needs and successful skills teaching should surely be included in all preparation programs. Assessment of these goals is elusive. Many programs require a clinical practicum, and this could serve as an assessment setting. However, this typically comes near the end of a program sequence, after the student has completed basic courses and courses in diagnosis and

remediation. Then, too, not all practica require a demonstration of the competencies listed above for successful course completion. Further, a single successful experience in a tutorial situation may not be adequate evidence that the teacher can indeed identify and remediate skills difficulties generally. These vagaries led to the development of an alternative technique for assessing these aspects of teaching competence.

Historical Development. The development of a series of assessment techniques for these aspects of teaching competence has had a common element. The texts, Systematic Reading Instruction (1972) and How to Teach Reading Systematically (1973) by Duffy and Sherman have offered the instructional framework from which the lessons were developed and evaluated. Initially students were required to submit written instructional models for teaching specific skills as well as a pretest-posttest designed to determine whether the subject(s) had mastered the skill. These written models were evaluated and returned either as passing or for revision and resubmitting (Duffy, 1971).

Recognizing the distance between developing a written model and classroom application, other strategies were developed. The second assessment technique was role-playing. The graduate student/teachers were required to develop a working model of a skills lesson and then present it to another graduate student/teacher who was to simulate an elementary school pupil. This technique seemed to narrow the distance between knowledge and application, but it was not without its problems. The student role-playing the elementary pupil often reacted quite unlike reality.

Throughout these stages the idea of a field assessment was constantly offered. Unfortunately the elementary pupils could not be brought in and there was neither the staff nor the money to accomplish field assessment. The concept of cassette evaluations evolved naturally from attempts to employ video taped assessment. Video taping, too, was not easily accomplished and, thus, audio-tapes were adopted as a logical alternative. This study, then, is an attempt to examine the relative effectiveness of evaluation of cassette mini-lessons in a micro-teaching setting.

Method

Subjects. Thirty graduate students enrolled in ERDG 606, Corrective Reading in the Classroom, at the State University of New York at Albany served as the population of the study. Each student was certified to teach in New York State; there were varying degrees of teaching experience and coursework in reading methods. Each graduate student was required to submit two audio-cassette tapes for evaluation, and to complete a mid-term and final examination in the paper-pencil multiple choice format.

Procedure. Students were provided with an instructional model for skills teaching. The model is presented in How to Teach Reading Systematically (Duffy and Sherman, et al, 1973). Following a verbal explanation of the instructional model, students were presented with demonstration lessons by the instructor. This was followed by the presentation of audio-taped models which were critiqued by the instructor with class

participation. Additional audio-models were available for students who wished to utilize them.

In addition to the instruction devoted to teaching skills lessons, students were provided information and models for assessing specific skills needs. Various instruments were offered and students were encouraged to use and/or develop informal techniques for assessing skills difficulties.

Students in the graduate course were generally employed as classroom teachers. Subjects for their skills mini-lessons were drawn from their classes. Students not employed as classroom teachers were required to participate in either a classroom or tutorial based field experience and drew their subjects from these populations. The objectives that were taught to were selected from Systematic Reading Instruction. Students pretested subjects until an appropriate objective had been identified. To verify the need for instruction the administration of the pretest was to be recorded at the beginning of the taped lesson and copies of the completed tests were to be submitted with the tape for evaluation. Students worked in either a tutorial or small(3-5)group situation for the taped mini lessons.

Two forms were constructed for assessment of the teaching tapes. The first was a checklist (see attachment 1) adapted from Appendix H of Systematic Reading Instruction by Duffy and Sherman (1972). This instrument had one column for student evaluation of the taped performance. The checklist included items on specific aspects of the teaching model; directing visual or auditory attention appropriately, modelling of activity,

reinforcement ratio and effectiveness, length of lesson relative to learner attention span, conciseness of presentation, clarity of directions, appropriateness of final student response in relation to initial objective,etc. The instructor rated the student on the same criteria and form.

The second form was developed for written instructor comments concerning specific portions of the teaching lesson. This form was broken into four parts: suitability of pre- and post-test forms, directing attention, presentation of lesson, and practice activity and/or materials. Thus, the instructor provided a written critique of each taped lesson as well as a completed checklist for comparison with students personal evaluation.

These assessment instruments were seen as formative. The taped instructional episodes were ranked on a 1 to 5 scale; in .5 increments, with 5 being the highest ranking attainable. The written comments provided both a critique and suggestions for improvement. Subjects submitting taped lessons rated at or below 2 were required to discuss the assessment with the instructor and resubmit the assignment. These resubmitted tapes were not included in the data analysis.

The mid-term and final examinations followed the traditional multiple choice format. Questions primarily tested cognitive knowledge about various facets of reading theory and instruction as presented in the texts and lectures. These tests were computer scored and scores ranged from 88 out of 100 to 58 out of the same. All items on the tests had been used with other similar

populations at both State University of New York at Albany and Michigan State University. Each question had thus been submitted to an item analysis inspection and refinement based upon these inspections.

Results. In determining the results of student performance on examinations and tapes, scores and ratings were simply averaged for each subject. The Kendall Tan was calculated for the averaged test scores and performance on the taped mini-lessons as rated (0-5) and ranked.

Test averages ranged from 88% to 58%. Averaged performance on taped mini-lessons ranged from 4.50 to 1.75. Calculation of T yielded .21 indicating there was no significant ($p < .05$) correlation between the two variables (see Figure 1).

INSERT FIGURE 1 ABOUT HERE

Discussion. Caution is advised in any attempt to interpret these results. While these data indicate no significant correlation exists between teacher performance on a traditional style multiple choice assessment instrument and teaching performance in a micro-teaching situation, one cannot yet concomitantly assume a high positive correlation between a micro-teaching performance and total classroom performance. While Allen (1966) argues this correlation exists, further examination of this comparison is a next logical step in the design and refinement of appropriate instruments and techniques for assessing teaching competence in teacher preparation programs. However, these

results conform to earlier studies which found little correlation between achievement on various paper/pencil assessment instruments and teaching ability thus, suggesting the audio-tapes are preferable assessment techniques. Even though only a singular aspect of teaching performance was measured, the data suggest support for the hypothesis that an audio-taped micro-teaching lesson is a suitable alternative to either traditional or field assessment. Finally, the data suggest that the widespread use of mastery tests in CBTE curricula should be seriously questioned. Especially if the tests are used in an attempt to predict or ensure a minimal level of teaching effectiveness.

Several comments on the practicality of the use of the audio-tapes must be made. First, while less time consuming than field assessment, the assessment of tapes is still quite a bit more rigorous and demanding than correcting paper/pencil examinations. Tapes cannot be skimmed, and our students seemed to have difficulty in limiting themselves to the suggested 12 minute limit (some extending beyond the 30 minute span). Second, from two years of experimentation several refinements in the technique have emerged. The checklists and instructor response forms are but two examples. These were necessitated by students who were brash enough to ask what specifically was lacking from taped lessons rated inadequate by an instructor. Unlike traditional examinations, the original tape evaluations offered no item by item critique, but rather a general critique and a subsequent adequate/inadequate rating. The legitimate demands of

some more vociferous students led to the use and modification of the Duffy and Sherman checklist and to the development of the five section instructor response form.

To ensure the taped mini-lessons did not become summative evaluations, all students receiving an inadequate rating are required to resubmit lessons until a minimum number of adequate lessons are received. Conferences with the instructor were encouraged and a number of model tapes, with critiques, were available for student use. Teaching, primarily through modelling, has to precede the assignment of micro-teaching lessons.

Conclusions. Audio-tapes supply the teacher trainer with another useful technique for assessment. They are convenient, easy to handle, relatively inexpensive, widely available, and require few technical skills for preparation. However, assessment of taped lessons is more time consuming than computer scored examinations. The teacher trainer can expect to spend at least 30 minutes on most taped lessons. Finally, teaching tapes require the teacher trainer to have numerous teaching skills and a firm grasp of instructional pedagogy. Through experience we have found that the use of teaching tapes requires explicit evaluative comments. Students want to know why they received a low ranking and they want specifics. Suggestions for improving the teaching lesson are typically anticipated even from those receiving high rankings. Thus, the teacher-trainer must necessarily be an instructional specialist, otherwise formative assessment is difficult.

In summary, audio-taped micro-teaching lessons seem to be viable alternatives to the traditional paper/pencil instruments. The taped assignments are more work and require an emphasis on teaching rather than an emphasis on storing cognitive knowledge. They do, however, seem to be a more appropriate assessment technique for teacher education programs which seek to ensure a minimal level of teaching competence in their graduates.

Bibliography

Allen, Dwight. "A New Design for Teacher Education: The Teacher Intern Program at Stanford University," Journal of Teacher Education, 17, No. 3 (Fall, 1966), pp. 296-300.

Allington, Richard. "A Rationale for the Competency-Based Education of Reading Teachers," Journal of Reading, 17, No. 7 (April 1974), pp.

Barr, A.S. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations," Journal of Experimental Education, 30 (September, 1961), pp. 1-166.

Dodl, Norman. "Program Evaluation and Pupil Assessment," in Cooper, DeVault, et al (eds.) Competency Based Teacher Education. (Berkely:McCutchan, 1973) p. 113.

Duffy, Gerald. "The Continued Development of Two Undergraduate Reading Methods Courses: A final report to the Educational Development Program, 1971-72," mimeographed paper from Department of Elementary and Special Education, Michigan State University, East Lansing, Michigan (July 1971).

Duffy, Gerald and Sherman, George. Systematic Reading Instruction, (New York: Harper-Row, 1972).

Duffy, Gerald; Sherman, George; Allington, R.; McElwee, M; Roehler, L. How to Teaching Reading Systematically, (New York: Harper-Row, 1973).

Getzels, Jacob, and Jackson, Phillip. "The Teacher's Personality and Characteristics," in Gage, Handbook of Research on Teaching, (Chicago: Rand-McNally, 1963), pp. 506-582.

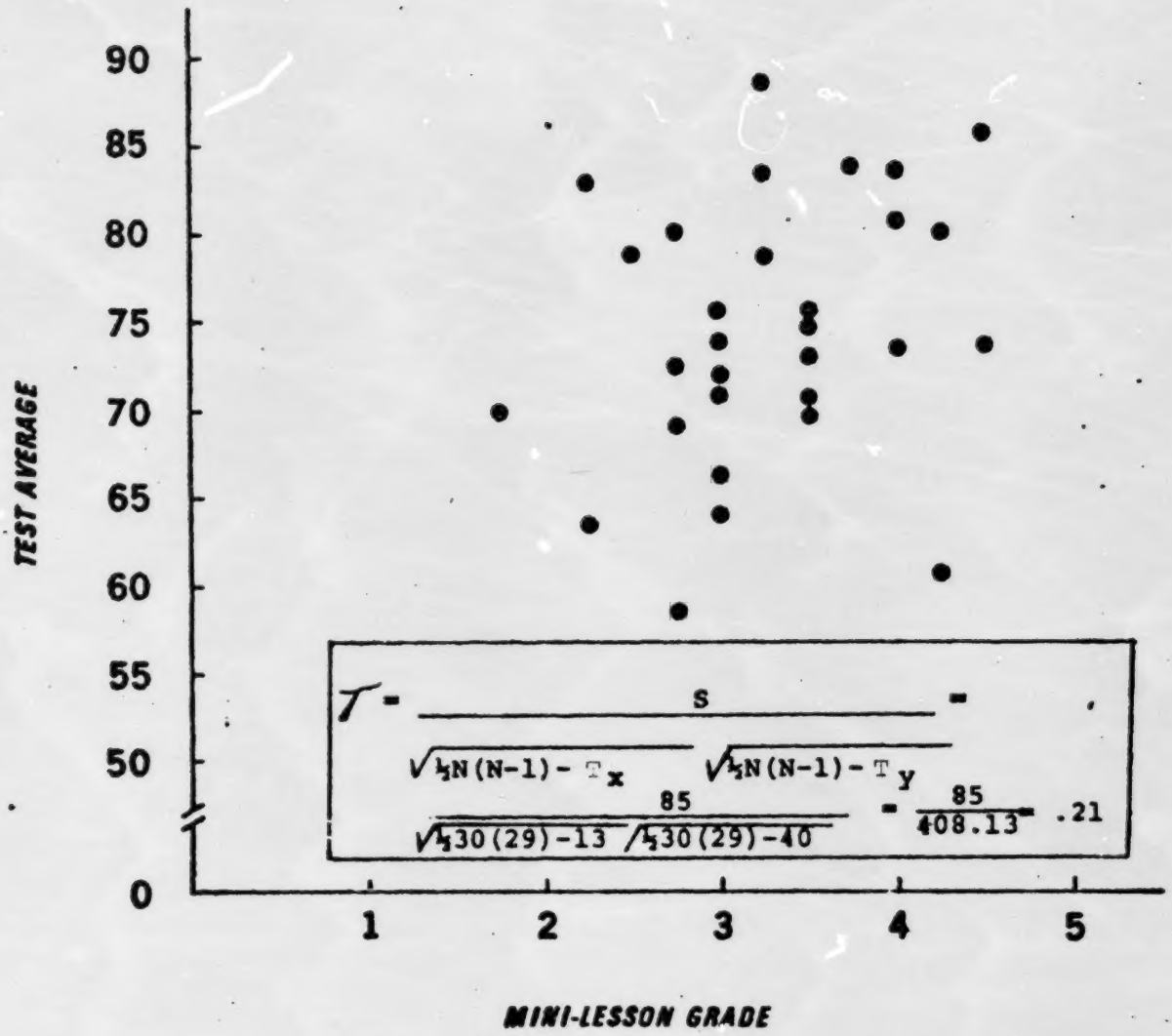
Levin, Henry M. "Recruiting Teachers for Large City School," in Studies in Social Economics. Brookings Institute, 1968.

Popham, James. "Performance Tests of Teaching Proficiency: Rationale, Development and Validation," American Educational Research Journal, 8 (January, 1971) pp. 105-117.

Shea, Joan. "The Predictive Value of Various Combinations of Standardized Tests and Subtests for Prognosis of Teaching Efficiency," Educational Research Monograph, 19, No. 6 (Washington, D.C.: Catholic University of America, 1955).

Walberg, H.J. "Scholastic Aptitude, the National Teacher Examinations, and Teaching Success," Journal of Educational Research, 61, No. 3 (November, 1967) pp. 129-131.

FIGURE 1


90
85
80
75
70
65
60
55
50
0
1
2
3
4
5
TEST AVERAGE
MINI-LESSON GRADE
$T = \frac{S}{\sqrt{\frac{1}{2}N(N-1) - T_x}\ \sqrt{\frac{1}{2}N(N-1) - T_y}} = \frac{85}{\sqrt{\frac{1}{2}30(29) - 13}\ \sqrt{\frac{1}{2}30(29) - 40}} = \frac{85}{408.13} = .21$

STATE UNIVERSITY OF NEW YORK AT ALBANY
Checklist for Evaluating Skill Lessons
ERDG 606

Skill taught _____ p. _____ Name _____

behavior	your evaluation		instructor evaluation	
A. Directing attention:	Yes	No	Yes	No
1. Have you been specific in directing attention?				
2. Have you directed attention appropriately?				
3. Have you used physical cues to direct attention?				
4. Have you established a good psychological set?				
B. Presentation:				
1. Is the presentation clear and concise?				
2. Is the lesson contained within the attention span of the learner?				
3. Did you proceed from simple to complex?				
4. Have you done less talking than the learner?				
5. Did you model the task for the learner?				
6. Have you highlighted the "pieces of the puzzle"?				
7. Have you provided diminishing crutches?				
8. Has the learner been lead to make an independent response?				
C. Response and Reinforcement:				
1. Has the learner had opportunities to respond?				
2. Does the learner do, as well as listen?				
3. Have you provided feedback on pupil responses?				
4. Have you provided positive reinforcement?				
5. Is the reinforcement effective?				

behavior	your evaluation		instructor evaluation	
D. Practice and Application:	Yes	No	Yes	No
1. Is the learner practicing the skill taught?				
2. Is the practice relevant? (as opposed to busy work)?				
3. Are instructions clear and comprehensible?				
4. Does the learner receive immediate feedback?				
5. Have I structured an experience to ensure transfer?				
Additional comments:				

State University of New York at Albany
Instructor Comment Sheet: Skill Lessons
ERDG 606

pre-test:

attender: physical -

psychological -

presentation:

practice and application:

evaluation _____

RD-16-1-A274